

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

On time consistency for mean-variance portfolio selection

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1753729> since 2021-02-09T14:13:27Z

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

On time consistency for mean-variance portfolio selection

Elena Vigna*

June 12, 2020

Abstract

This paper addresses a comparison between different approaches to time inconsistency for the mean-variance portfolio selection problem. We define a suitable intertemporal preferences-driven reward and use it to compare three common approaches to time inconsistency for the mean-variance portfolio selection problem over $[t_0, T]$: precommitment approach, consistent planning or game theoretical approach, and dynamically optimal approach. We prove that, while the precommitment strategy beats the other two strategies (that is a well-known obvious result), the consistent planning strategy dominates the dynamically optimal strategy until a time point $t^* \in (t_0, T)$ and is dominated by the dynamically optimal strategy from t^* onwards. Existence and uniqueness of the break even point t^* is proven.

Keywords. Time inconsistency, dynamic programming, Bellman's optimality principle, precommitment approach, consistent planning approach, mean-variance portfolio selection.

JEL classification: C61, D81, G11.

*University of Torino and Collegio Carlo Alberto, Italy. Corso Unione Sovietica 218 bis, 10134, Torino, Italy. E-Mail: elena.vigna@unito.it. Tel. +39 011 670 5754.

1 Introduction and review of the literature

The notion of time inconsistency for optimization problems dates back to Strotz (1956). Broadly speaking, time inconsistency arises in an intertemporal optimization problem when the optimal strategy selected at some time t is no longer optimal at time $s > t$. In other words, a strategy is time-inconsistent when the individual at future time $s > t$ is tempted to deviate from the strategy decided at time t . It is well known that an optimization problem gives rise to time-inconsistent strategies when the Bellman's principle does not hold and dynamic programming cannot be applied. In finance, a notable example of problem which is time-inconsistent is the mean-variance problem, where the time inconsistency is due to the fact that there is a non-linear function of the expectation of final wealth in the optimization criterion (due to the presence of the variance of final wealth). Another important problem which produces a time-inconsistent behaviour is the investment-consumption problem with non-exponential discounting. This was the case studied by Strotz (1956), and time inconsistency arises because the initial point in time enters in an essential way the objective criterion. A third major type of time inconsistency is due to the initial/present state entering the investor's utility function, see for example Christensen & Lindensjö (2018). For a clarifying formalization of the possible sources of time inconsistency in intertemporal optimization problems, see Björk & Murgoci (2014) for the discrete time framework and Björk, Khapko & Murgoci (2017) for the continuous-time framework.

In the last two decades there has been a renewed interest in time inconsistency for financial and economic problems. According to Strotz (1956) there are two possible ways to deal with time-inconsistent problems: (i) precommitment approach; (ii) consistent planning or game theoretical approach. In the precommitment approach, the controller fixes an initial point (t_0, x_0) and finds the optimal control law \hat{u} that maximizes the objective functional at time t_0 with wealth x_0 , $J(t_0, x_0, u)$, disregarding the fact that at future time $t > t_0$ the control law \hat{u} will *not* be the maximizer of the objective functional at time t with wealth x , $J(t, x, u)$; therefore, he precommits to follow the initial strategy \hat{u} , despite the fact that at future dates he will no longer be optimal according to his criterion. To do this, he must be able to precommit his future selves to the strategy decided at time t_0 . In the consistent planning approach, one tries to avoid time inconsistency by selecting the *“best plan among those that*

he will actually follow” (Strotz, 1956). This approach translates into the search of a Nash subgame perfect equilibrium point. Intuitively, sitting at time t the future time interval $[t, T]$ can be seen as a continuum of players, each player $s \geq t$ being the “reincarnation” at time s of the player who sits at time t . With this approach, a time-consistent equilibrium policy is the collection of all decisions $\hat{u}(s, \cdot)$ taken by any player $s \in [t, T]$, such that if player t knows that all players coming after him (in (t, T)) will use the control \hat{u} , then it is optimal to him, too, to play control \hat{u} .

The literature is full of examples of applications of the two approaches outlined. For conciseness reasons, we here report only a few of them. For instance, the mean-variance portfolio selection problem has been solved with the precommitment approach by Richardson (1989), Bajeux-Besnainou & Portait (1998), Zhou & Li (2000) and Li & Ng (2000), the first two with the martingale method, the last two with an embedding technique that transforms the mean-variance problem into a standard linear-quadratic control problem. The game theoretical solution to the mean-variance problem has been found originally by Basak & Chabakauri (2010), then extended to a more general class of time-inconsistent problems by Björk et al. (2017). Other papers on the consistent planning approach for the mean-variance problem are Björk, Murgoci & Zhou (2014), Czichowsky (2013).

The precommitment strategy and the game theoretical approach are not the only ways to attack a problem that gives rise to time inconsistency. An alternative approach has been introduced by Pedersen & Peskir (2017) for the mean-variance portfolio selection problem, namely, the dynamically optimal strategy. This approach is novel to treat time inconsistency, although related work can be found in Karnam, Ma & Zhang (2017). The strategy proposed by Pedersen & Peskir (2017) is time-consistent in the sense that it does not depend on initial time and initial state variable, but differs from the subgame perfect equilibrium strategy. Moreover, their policy is intuitive and formalizes a quite natural approach to time inconsistency: it represents the behaviour of an optimizer who continuously reevaluates his position and solves infinitely many problems in an instantaneously optimal way. The dynamically optimal individual is similar to the continuous version of the naive individual described by Pollak (1968). At each time t the dynamically optimal investor turns out to be the “reincarnation” of the precommitted investor, for at time t he plays the strategy that the time- t precommitted investor would play, forgetting about his past and ignoring his future,

and deviates from it immediately after, by wearing the clothes of the time t^+ precommitted investor. It is worth noting that the dynamically optimal approach has strong similarities also with the receding horizon procedure or the model predictive control — so-called rolling horizon procedures, see Powell (2011) — that are well established methods of repeated optimization over a rolling horizon for engineering optimization problems with an infinite time horizon.¹

It is reasonable to assume that which of the three approaches just described should be used in a time-inconsistent optimization problem depends on the investor and on the nature the specific economic phenomenon under consideration. This paper adds to the debate on what is the “appropriate” approach to time inconsistency by investigating the differences among the three approaches described above. We define a suitable intertemporal preferences-driven reward and use it to compare the three outlined approaches to time inconsistency for the mean-variance portfolio selection problem: precommitment, consistent planning and dynamic optimality. The idea behind this intertemporal preferences-driven reward is simple: the reward related to future time t (as evaluated at time t_0) associated to a given approach u is the expectation at initial time t_0 of the performance criterion experienced at future time t using the approach u . A preview of the results is the following. Expectedly and obviously, the precommitment strategy beats the other strategies if the investor is able to precommit, only cares at the view point at time t_0 and is not concerned to be time-inconsistent in (t_0, T) ; if instead he is not able to precommit he should go for one of the two time-consistent strategies: in this case we find that the Nash-equilibrium strategy dominates the dynamically optimal strategy until a time point $t^* \in (t_0, T)$ and is dominated by the dynamically optimal strategy from t^* onwards. We prove existence and uniqueness of the break even point t^* and provide a closed form for it. Interestingly, the break even point t^* does not depend on wealth and risk aversion, while it increases with the market price of risk and the time horizon T . These results are in line with the results in Pedersen & Peskir (2017), who also address the comparison among the two time-consistent strategies. Differently from them, we make the comparison at any time $t \in [t_0, T]$ and not only at initial time t_0 and final time T .

This is not the only paper on the comparison among these three approaches to time in-

¹In the problem considered in this paper the time interval over which the optimization is done shrinks when time passes, while it remains fixed in the problems where the receding horizon procedure is applied.

consistency. In a companion paper, Vigna (2017) disentangles the notion of time consistency into the two notions of tail optimality of a control map and preferences consistency of an optimizer, in the attempt to shed light on the differences between time-consistent and time-inconsistent optimization problems, and between the three approaches to time inconsistency. As for the comparison between two of the three approaches to time inconsistency, a comparison between the naive and the sophisticated approaches is in Chen, Li & Zeng (2014), who study the optimal dividend problem of an insurance company in the presence of time inconsistency generated by non-exponential discount factor; Cong & Oosterlee (2016) establish a link between the time-consistent and the precommitment investment strategies in a defined contribution pension scheme; Wang & Forsyth (2011) compare time-consistent versus precommitment policies and compare their corresponding efficient frontiers for a mean-variance optimization problem; Van Staden, Dang & Forsyth (2018) compare the precommitment and the time-consistent strategies in the presence of realistic investment constraints; Cui, Li & Shi (2017) highlight the drawbacks of precommitment and game theoretical strategies, and investigate a self-coordination policy that aims at balancing global interest and local interests of the decision-maker; Bensoussan, Wong & Yam (2019) analyze the effect of constraints on the value function of both precommitment and game theoretical approaches, and find the unexpected result that for the game theoretical approach the presence of constraints can improve the payoff, while for the precommitment approach this paradox does not occur; Menoncin & Vigna (2020) compare the precommitment and the dynamically optimal strategies for a defined contribution pension scheme and find that the dynamically optimal strategy reacts better to extreme scenarios of market returns (either good or bad) because of the continuous adjustment of the final target. Two papers that attack with a different angle the time inconsistency of the precommitment strategy are Cui, Li, Wang & Zhu (2012) and Shi, Li & Cui (2017), who propose a weaker notion of time consistency (neither related to the game theoretical nor to the dynamically optimal approach) and find that, while the precommitment strategy in the continuous time in the presence of a pure diffusive market satisfies it, the same is not true for the multi-period time or in a jump diffusion market; they propose semi-self-financing revised mean-variance strategies that beat the precommitment one.

The remainder of the paper is as follows. In Section 2, we formulate the mean-variance

optimization problem, specify the financial market and list the strategies corresponding to each of the three approaches to time inconsistency. In Section 3, we define the intertemporal preferences-driven reward and compare the three strategies according to it at all times between initial time and final time. Section 4 concludes.

2 The mean-variance portfolio selection problem

2.1 Statement of the problem

An investor, endowed with a wealth $x_0 > 0$ at time $t_0 \geq 0$, is faced with a portfolio selection problem on the time horizon $[t_0, T]$, $T \geq t_0$. The financial market available for the portfolio allocation problem is the Black-Scholes model (see e.g. Björk, 1998). This consists of two assets, a riskless one, whose price $B(t)$ follows the dynamics:

$$dB(t) = rB(t)dt,$$

where $r > 0$, and a risky asset, whose price dynamics $S(t)$ follows a geometric Brownian motion with drift $\lambda \geq r$ and volatility $\sigma > 0$:

$$dS(t) = \lambda S(t)dt + \sigma S(t)dW(t),$$

where $W(t)$ is a standard Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t \in [t_0, T]}, \mathbb{P})$, with $\mathcal{F}_t = \sigma\{W(s) : s \leq t\}$ the natural filtration. The amount of money invested in the risky asset at time t is denoted by $u(t)$. The fund at time t under control u , $X^u(t)$, grows according to the following SDE:

$$\begin{aligned} dX^u(t) &= [u(t)(\lambda - r) + rX^u(t)]dt + u(t)\sigma dW(t), \\ X^u(t_0) &= x_0 > 0. \end{aligned} \tag{1}$$

The investor is a mean-variance optimizer and his aim is to solve the problem

$$\sup_{u \in \mathcal{U}} J(t_0, x_0, u) \equiv [\mathbb{E}_{t_0, x_0}(X^u(T)) - \alpha \mathbb{V}_{t_0, x_0}(X^u(T))], \tag{2}$$

where $\alpha > 0$ is a measure of risk aversion, and \mathcal{U} is the set of admissible strategies, defined as the set of \mathbb{R} -valued stochastic processes $u = \{u(t)\}_{t \in [t_0, T]}$ that are \mathcal{F}_t -adapted and s.t. the SDE (1) has a unique strong solution.

2.2 Three possible approaches

Because Problem (2) is time-inconsistent, the investor can adopt three alternative investment strategies, that are optimal according to different perspectives. The three approaches are: (i) precommitment (ii) consistent planning or Nash equilibrium (iii) dynamical optimality.

Precommitment approach

With the precommitment approach, one fixes the initial point (t_0, x_0) and finds, if it exists, the control law u^p that maximizes only $J(t_0, x_0, u)$, i.e., the precommitment strategy. This is formalized by the following definition.

Definition 2.1. *Given the mean-variance problem (2), the strategy u^p that maximizes $J(t_0, x_0, u)$, i.e., the control map*

$$u_{t_0, x_0}^p : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R} \quad (3)$$

such that

$$J(t_0, x_0, u^p) = \sup_{u \in \mathcal{U}} J(t_0, x_0, u)$$

if it exists, is called the precommitment strategy.

Consistent planning, Nash equilibrium approach

According to the consistent planning approach, in order to solve the mean-variance problem (2), one should choose “the best plan among those that he will actually follow”. The construction of this strategy is based on the game theoretic interpretation: (i) consider a non-cooperative game where to each point in time t is associated Player t ; (ii) Player t chooses the control at time t with the control function $u(t, \cdot)$, so that if the value of the wealth at time t is X_t Player t plays $u(t, X_t)$; the collection of all the control functions of all players for $t \in [t_0, T]$ gives a feedback control law $u : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$. The feedback control

law u^e is a subgame perfect Nash equilibrium if for each t the following property holds: if for each $s > t$ Player s plays $u^e(s, \cdot)$ then it is optimal for Player t too to play $u^e(t, \cdot)$.

More rigorously, the definition of subgame perfect equilibrium for the continuous time case is as follows (see Björk et al., 2017):

Definition 2.2. *Consider an admissible control law u^e . Choose an arbitrary admissible control law $u \in \mathcal{U}$ and a fixed real number $h > 0$. Fix an arbitrarily chosen initial point (t, x) and define the control law u_h by*

$$u_h(s, y) = \begin{cases} u(s, y) & \text{for } t \leq s < t + h, y \in \mathbb{R} \\ u^e(s, y) & \text{for } t + h \leq s \leq T, y \in \mathbb{R} \end{cases} \quad (4)$$

If

$$\liminf_{h \rightarrow 0} \frac{J(t, x, u^e) - J(t, x, u_h)}{h} \geq 0$$

for all $u \in \mathcal{U}$, then we say that u^e is an equilibrium strategy.

The equilibrium strategy for the mean-variance problem is found by solving an extended Hamilton-Jacobi-Bellman equation for the value function, see Björk & Murgoci (2010), Proposition 4.1 and Theorem 4.1.

Dynamically optimal approach

We illustrate the construction of the dynamically optimal strategy introduced by Pedersen & Peskir (2017) for a the mean-variance problem (2) in 3 steps.

Step 1. Given the mean-variance problem (2), assume that for the fixed initial point (t_0, x_0) the precommitment strategy maximizing the criterion $J(t_0, x_0, u)$ exists² and is given by (see (3)):

$$u_{t_0, x_0}^p : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}. \quad (5)$$

Step 2. Define the new control map

$$u^d(s, y) = u_{s, y}^p(s, y), \quad \text{for } (s, y) \in [t_0, T] \times \mathbb{R}, \quad (6)$$

²It is worth noting that Pedersen & Peskir (2017) proved existence of the static precommitment strategy.

where the right hand side of (6) is obtained by replacing t_0 with s and x_0 with y in the function (5).

Step 3. The strategy $u^d(s, y) : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, is called the *dynamically optimal strategy*.

2.3 The three strategies

If the investor cares *only* at being optimal at time t_0 , is able to precommit his behaviour in the future, and does not care of being time-inconsistent at future time $t > t_0$, then he could adopt the precommitment strategy u^p (see Zhou & Li, 2000):

$$1. \text{ precommitment policy} \quad u_{t_0, x_0}^p(t, x) = \frac{\delta}{\sigma} \left[x_0 e^{r(t-t_0)} - x + \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)} \right], \quad (7)$$

where $\delta = (\lambda - r)/\sigma$ is the market price of risk, or Sharpe ratio.

If the investor is not able to precommit and cares *also* at being time-consistent at future time $t > t_0$, then he can adopt either the time-consistent Nash equilibrium policy u^e (see Basak & Chabakauri, 2010, and Björk et al., 2014):

$$2. \text{ Nash equilibrium policy} \quad u_{t_0, x_0}^e(t, x) = \frac{\delta}{\sigma} \frac{1}{2\alpha} e^{-r(T-t)}, \quad (8)$$

or the time-consistent dynamically optimal policy u^d (see Pedersen & Peskir, 2017):

$$3. \text{ dynamically optimal policy} \quad u_{t_0, x_0}^d(t, x) = \frac{\delta}{\sigma} \frac{1}{2\alpha} e^{(\delta^2-r)(T-t)}. \quad (9)$$

If the investor selects the precommitment policy, his optimal wealth follows the dynamics:

$$X_t^p = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2-r)(T-t)} \left[e^{\delta^2(t-t_0)} - e^{-\delta(W_t - W_{t_0}) - \frac{\delta^2}{2}(t-t_0)} \right]; \quad (10)$$

if he selects the Nash-equilibrium policy, his optimal wealth follows the dynamics:

$$X_t^e = x_0 e^{r(t-t_0)} + \frac{\delta}{2\alpha} e^{-r(T-t)} [\delta(t - t_0) + W_t - W_{t_0}]; \quad (11)$$

if he selects the dynamically optimal policy, his optimal wealth follows the dynamics:

$$X_t^d = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2-r)(T-t)} \left[e^{\delta^2(t-t_0)} - 1 + \delta \int_{t_0}^t e^{\delta^2(t-s)} dW_s \right]. \quad (12)$$

In all cases, the value function associated to the adoption of strategy u is

$$\begin{aligned} V^u : [t_0, T] \times \mathbb{R} &\rightarrow \mathbb{R} \\ V^u(t, x) &= \mathbb{E}_{t,x}(X^u(T)) - \alpha \mathbb{V}_{t,x}(X^u(T)). \end{aligned} \quad (13)$$

3 Comparison among different strategies

In this section, we show how to make a comparison among the three approaches illustrated in Section 2.2.

We shall start with a brief inspection of the distribution of final wealth in the three approaches. Using (10), (11) and (12), we have:

$$X_T^p = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} \left[e^{\delta^2(T-t_0)} - e^{-\delta(W_T - W_{t_0}) - \frac{\delta^2}{2}(T-t_0)} \right],$$

$$X_T^e = x_0 e^{r(T-t_0)} + \frac{\delta}{2\alpha} [\delta(T-t_0) + W_T - W_{t_0}],$$

and

$$X_T^d = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} \left[e^{\delta^2(T-t_0)} - 1 + \delta \int_{t_0}^T e^{\delta^2(T-s)} dW_s \right].$$

Straightforward calculations yield

$$\mathbb{E}_{t_0, x_0}(X_T^p) = \mathbb{E}_{t_0, x_0}(X_T^d) = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} [e^{\delta^2(T-t_0)} - 1]$$

while

$$\mathbb{E}_{t_0, x_0}(X_T^e) = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} [\delta^2(T-t_0)]$$

Because

$$e^{\delta^2(T-t_0)} - 1 \geq \delta^2(T-t_0),$$

we see that the expected final wealth of precommitment and dynamically optimal strategies coincide and are larger than that of the Nash equilibrium strategy.

However, in order to provide a complete picture of the comparison, we need to incorporate also the mean-variance criterion and we need to perform the comparison at different points in time. The comparison will be performed both at initial time t_0 and also at future time $t > t_0$, by defining proper stochastic reward functions at time t for each strategy.

In order to make a comparison of the three strategies, we will henceforth imagine that we have three investors: the precommitted static one (P-investor), the Nash-equilibrium one (E-investor) and the dynamically optimal one (D-investor), who will adopt the policies (7), (8) and (9), respectively.

3.1 Comparison among strategies at initial time t_0

The comparison among the three strategies at initial time t_0 is straightforward. It suffices to compare the three value functions at t_0 .

Adopting the precommitment strategy (7), the value function at time t_0 with wealth x_0 is

$$V^{u^p}(t_0, x_0) = \mathbb{E}_{t_0, x_0}(X^{u^p}(T)) - \alpha \mathbb{V}_{t_0, x_0}(X^{u^p}(T)) = x_0 e^{r(T-t_0)} + \frac{e^{\delta^2(T-t_0)} - 1}{4\alpha}. \quad (14)$$

Adopting the Nash equilibrium strategy (8), the value function at time t_0 with wealth x_0 is

$$V^{u^e}(t_0, x_0) = \mathbb{E}_{t_0, x_0}(X^{u^e}(T)) - \alpha \mathbb{V}_{t_0, x_0}(X^{u^e}(T)) = x_0 e^{r(T-t_0)} + \frac{\delta^2(T-t_0)}{4\alpha}. \quad (15)$$

Adopting the dynamically optimal strategy (9), the value function at time t_0 with wealth x_0 is

$$V^{u^d}(t_0, x_0) = \mathbb{E}_{t_0, x_0}(X^{u^d}(T)) - \alpha \mathbb{V}_{t_0, x_0}(X^{u^d}(T)) = x_0 e^{r(T-t_0)} + \frac{4e^{\delta^2(T-t_0)} - e^{2\delta^2(T-t_0)} - 3}{8\alpha}. \quad (16)$$

The following results hold:

Proposition 3.1. *For all $t_0 \leq T$ and $x_0 > 0$*

$$V^{u^p}(t_0, x_0) \geq V^{u^e}(t_0, x_0) \geq V^{u^d}(t_0, x_0). \quad (17)$$

The equalities hold if and only if $t_0 = T$, or $\delta = 0$.

Proof. For the first of the two inequalities:

$$V^{u^p}(t_0, x_0) - V^{u^e}(t_0, x_0) = \frac{e^{\delta^2(T-t_0)} - 1 - \delta^2(T-t_0)}{4\alpha} \geq 0,$$

where the inequality holds as an equality if and only if $\delta^2(T-t_0) = 0$, that holds if and only if $t_0 = T$ or $\delta = 0$.

For the second of the two inequalities:

$$V^{u^e}(t_0, x_0) - V^{u^d}(t_0, x_0) = \frac{e^{2\delta^2(T-t_0)} + 2\delta^2(T-t_0) + 3 - 4e^{\delta^2(T-t_0)}}{8\alpha} \geq 0,$$

where the inequality is due to the facts that (i) $\delta^2(T-t_0) \geq 0$, (ii) the function $f(x) = e^{2x} + 2x + 3 - 4e^x$ is strictly increasing, and (iii) $f(0) = 0$. (ii) and (iii) imply that the inequality holds as an equality if and only if $t_0 = T$ or $\delta = 0$. \square

Remark 1. We see from Proposition 3.1 that, considering the reward only at time t_0 , the P-investor receives a higher value function than the E-investor and the D-investor. This is obvious, because the precommitment strategy by definition maximizes the objective criterion at initial time, see also Wang & Forsyth (2012). The second inequality, already found in Pedersen & Peskir (2017), shows that at initial time t_0 the Nash-equilibrium strategy provides a higher value function than the dynamically optimal strategy. This is also expected and consistent with the fact that, by construction, the Nash-equilibrium strategy is *the best* among all the time-consistent strategies (see Strotz, 1956).

3.2 Comparison among strategies at time t : Reward functions and expected reward functions

The comparison cannot be done only at time t_0 , otherwise the obvious answer to the question *what is the best strategy to be adopted* is “the precommitment strategy”, that beats all the others from the point of view at t_0 (provided that it will be actually followed also after t_0). Indeed, if the investor is able to precommit and only cares of being mean-variance at time t_0 , he will select the precommitment strategy, and will not care of being time-inconsistent after t_0 .

Suppose instead that the investor cannot precommit and is concerned of being time-consistent at every $t \in [t_0, T]$. Then, assuming that he will not change his mean-variance preferences and his risk aversion α , his criterion at every time $t \in [t_0, T]$ will still be to maximize the mean of final wealth while minimizing its variance. Therefore, it is reasonable to assume that the **reward for the mean-variance investor** at time t with wealth x adopting strategy u is:

$$J^u(t, x) = \mathbb{E}_{t,x}(X_T^u) - \alpha \mathbb{V}_{t,x}(X_T^u). \quad (18)$$

Let us notice that the comparison among the three investors at time $t > t_0$ is delicate, because, while at time t_0 they have the same wealth x_0 , at time $t > t_0$ they have different wealths, because they have been following three different investment strategies from t_0 to t . The P-investor will have wealth X_t^p given by (10), the E-investor will have wealth X_t^e given by (11), the D-investor will have wealth X_t^d given by (12), and in general these levels of wealth will be different from each other. Nevertheless, considering (18), their degree of happiness can be measured by their rewards:

$$J^p(t, X_t^p) = \mathbb{E}_{t,X_t^p}(X_T^p) - \alpha \mathbb{V}_{t,X_t^p}(X_T^p) \quad \text{reward for the P-investor at time } t$$

$$J^e(t, X_t^e) = \mathbb{E}_{t,X_t^e}(X_T^e) - \alpha \mathbb{V}_{t,X_t^e}(X_T^e) = V^e(t, X_t^e) \quad \text{reward for the E-investor at time } t$$

$$J^d(t, X_t^d) = \mathbb{E}_{t,X_t^d}(X_T^d) - \alpha \mathbb{V}_{t,X_t^d}(X_T^d) = V^d(t, X_t^d) \quad \text{reward for the D-investor at time } t.$$

We notice that in the last two cases, the reward at time t coincides with the value function

at time t , because the Nash-equilibrium and the dynamically optimal strategies are time-consistent. In the first case, the reward at time t does not coincide with the value function, because the investor follows the time-inconsistent strategy decided at time t_0 and therefore does not optimize the mean-variance criterion at any time $t > t_0$.

Notice that, standing at time t_0 , the rewards $J^p(t, X_t^p)$, $V^e(t, X_t^e)$ and $V^d(t, X_t^d)$ that refer to time $t > t_0$ are random variables. However, it is possible to compare them standing at time t_0 , by comparing their time- t_0 expectations. We thus define the following **expected reward functions**:

$$R(t, u^p; t_0, x_0) = \mathbb{E}_{t_0, x_0} (J^p(t, X_t^p)) \quad (19)$$

$$R(t, u^e; t_0, x_0) = \mathbb{E}_{t_0, x_0} (V^e(t, X_t^e)) \quad (20)$$

$$R(t, u^d; t_0, x_0) = \mathbb{E}_{t_0, x_0} (V^d(t, X_t^d)) . \quad (21)$$

In general, the expected value in t_0 of the reward at time t of the investor who follows the strategy u is

$$R(t, u; t_0, x_0) = \mathbb{E}_{t_0, x_0} [\mathbb{E}_{t, X_t^u} (X_T^u) - \alpha \mathbb{V}_{t, X_t^u} (X_T^u)] . \quad (22)$$

The motivation for this definition of expected reward function (on which we build the comparison among the strategies) is the following. If the investor preferences remain the same throughout the time period from t_0 to T , the investor's happiness will be measured at each future time t by the performance criterion $J^u(t, X_t^u)$ (u being the strategy adopted). Because the wealth of the three investors is the same only at time t_0 , the simplest way to compare the three stochastic processes $J^u(t, X_t^u)$ (where $u = p, e, d$) is to calculate their expected values at time t_0 as functions of time t , and then make a comparison between these functions at different (relevant) points in time $t \in [t_0, T]$.

Remark 2. The expected reward just introduced is probably the simplest way to deal with the different intertemporal rewards provided by the adoption of different strategies. It is not the only way. Another possible comparison among the three stochastic performance criteria and a way to introduce different weights to different points in time would be to calculate for

each strategy the quantity

$$\int_{t_0}^T w(t)R(t, u; t_0, x_0)dt \quad u = p, e, d$$

and then compare these three quantities. Clearly, the function $w(t)$ would reflect the subjective importance that the time- t_0 investor assigns to future time point t . We leave this possible comparison to future research.

3.2.1 Comparison of expected reward functions at times t_0 and T

For the comparison among the expected rewards of the three strategies at times t_0 and T , the following results hold:

Proposition 3.2. *If $t_0 < T$, $\delta \neq 0$, and $x_0 > 0$, then*

$$R(t_0, u^p; t_0, x_0) > R(t_0, u^e; t_0, x_0) > R(t_0, u^d; t_0, x_0) \quad (23)$$

and

$$R(T, u^p; t_0, x_0) = R(T, u^d; t_0, x_0) > R(T, u^e; t_0, x_0) \quad (24)$$

Proof. For every strategy u

$$R(t_0, u; t_0, x_0) = \mathbb{E}_{t_0, x_0} [\mathbb{E}_{t_0, x_0}(X_T^u) - \alpha \mathbb{V}_{t_0, x_0}(X_T^u)] = V^u(t_0, x_0).$$

Claim (23) follows by Proposition 3.1.

For every strategy u ,

$$\mathbb{E}_{T, X_T^u}(X_T^u) - \alpha \mathbb{V}_{T, X_T^u}(X_T^u) = X_T^u \quad \Rightarrow \quad R(T, u; t_0, x_0) = \mathbb{E}_{t_0, x_0}(X_T^u).$$

Using the dynamics (10), (11) and (12), we get:

$$\mathbb{E}_{t_0, x_0}(X_T^p) = \mathbb{E}_{t_0, x_0}(X_T^d) = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} \left(e^{\delta^2(T-t_0)} - 1 \right) \quad (25)$$

and

$$\mathbb{E}_{t_0, x_0}(X_T^e) = x_0 e^{r(T-t_0)} + \frac{\delta^2(T-t_0)}{2\alpha}.$$

If $\delta^2(T-t_0) \neq 0$, then $e^{\delta^2(T-t_0)} - 1 > \delta^2(T-t_0)$. Therefore, claim (24) is obtained. \square

Remark 3. Result (24) was already found by Pedersen & Peskir (2017) (Equation (4.4)), in a first attempt of comparison between the two time-consistent strategies for the mean-variance problem. In their comparative analysis, they did not provide detailed argumentation for the reasonableness of the method used for the comparison.

3.2.2 Comparison of expected reward functions at time t : comparison between time-consistent strategies

From Proposition 3.2, we see that the Nash-equilibrium strategy provides a higher reward than the dynamic strategy at initial time t_0 , and a lower reward than the dynamic strategy at final time T , suggesting the occurrence of a swap between the two strategies at some time $t^* \in (t_0, T)$. The next theorem shows the existence and uniqueness of such a break even point.

Theorem 3.3. *If $t_0 < T$, $\delta \neq 0$, and $x_0 > 0$, there exists one and only one point $t^* \in (t_0, T)$ such that*

$$R(t^*, u^e; t_0, x_0) = R(t^*, u^d; t_0, x_0). \quad (26)$$

The break even point t^ is the unique solution of the equation*

$$e^{2\delta^2(T-t)} - \left(4e^{\delta^2(T-t_0)} + 4\delta^2 t_0 - 3\right) + 2\delta^2(T+t) = 0. \quad (27)$$

Proof. By defining the function

$$\Delta R^{ed}(t) = R(t, u^e; t_0, x_0) - R(t, u^d; t_0, x_0), \quad (28)$$

claim (26) is equivalent to prove the existence and the uniqueness of a root of the function $\Delta R^{ed}(t)$ in the interval (t_0, T) . Proposition 3.2 yields:

$$\Delta R^{ed}(t_0) > 0 \quad \text{and} \quad \Delta R^{ed}(T) < 0. \quad (29)$$

Recalling that both the Nash-equilibrium strategy and the dynamic strategy are time-consistent, we can obtain $V^e(t, X_t^e)$ and $V^d(t, X_t^d)$ just by replacing t_0 with t , and x_0 with X_t^e and X_t^d in (15) and (16), respectively.

$$V^e(t, X_t^e) = X_t^e e^{r(T-t)} + \frac{\delta^2(T-t)}{4\alpha}. \quad (30)$$

$$V^d(t, X_t^d) = X_t^d e^{r(T-t)} + \frac{4e^{\delta^2(T-t)} - e^{2\delta^2(T-t)} - 3}{8\alpha}. \quad (31)$$

Using (20), (21), (30) and (31), we have

$$R(t, u^e; t_0, x_0) = \mathbb{E}_{t_0, x_0} (X_t^e) e^{r(T-t)} + \frac{\delta^2(T-t)}{4\alpha}, \quad (32)$$

and

$$R(t, u^d; t_0, x_0) = \mathbb{E}_{t_0, x_0} (X_t^d) e^{r(T-t)} + \frac{4e^{\delta^2(T-t)} - e^{2\delta^2(T-t)} - 3}{8\alpha}. \quad (33)$$

Using the closed-form expressions (11) and (12), we get

$$\mathbb{E}_{t_0, x_0} (X_t^e) = x_0 e^{r(t-t_0)} + \frac{\delta^2}{2\alpha} (t - t_0) e^{-r(T-t)}, \quad (34)$$

and

$$\mathbb{E}_{t_0, x_0} (X_t^d) = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2-r)(T-t)} \left[e^{\delta^2(t-t_0)} - 1 \right]. \quad (35)$$

By plugging (34) and (35) into (32) and (33), respectively, and using (28), after some simplifications we get

$$\Delta R^{ed}(t) = \frac{e^{2\delta^2(T-t)} + (3 - 4e^{\delta^2(T-t_0)} - 4\delta^2 t_0) + 2\delta^2(T+t)}{8\alpha}. \quad (36)$$

The function $\Delta R^{ed}(t)$ is continuous and, due to (29), it takes different signs at the extremes of $[t_0, T]$. Moreover,

$$\frac{d}{dt} (\Delta R^{ed}(t)) = -\frac{\delta^2}{4\alpha} (e^{2\delta^2(T-t)} - 1) < 0 \quad \text{for } t \in [t_0, T]$$

that implies that $\Delta R^{ed}(t)$ is also strictly decreasing. Therefore, there exists a unique root of $\Delta R^{ed}(t)$ in (t_0, T) and is given by the unique t^* that nullifies the numerator of (36). This concludes the proof. \square

Remark 4. Let us notice that Theorem 3.3 implies that

$$R(t, u^e; t_0, x_0) > R(t, u^d; t_0, x_0) \quad \forall t \in [t_0, t^*)$$

and

$$R(t, u^e; t_0, x_0) < R(t, u^d; t_0, x_0) \quad \forall t \in (t^*, T],$$

meaning that, among the two time-consistent strategies for the mean-variance problem, the Nash-equilibrium strategy provides on average a higher reward until time t^* , while the dynamic strategy provides on average a higher reward from time t^* onwards. This result is meaningful and suggests that the importance allocated by the decision-maker to different points in time should affect his attitude toward time inconsistency, and should play a part in the entire decision-making process. The first inequality is also consistent with the fact that the Nash equilibrium strategy is the *best* plan among those that are time-consistent (see Strotz, 1956): the criterion “best” here indicates the best at initial time t_0 , and indeed $R(t_0, u^e; t_0, x_0) > R(t_0, u^d; t_0, x_0)$. Clearly, selecting the best time-consistent plan only from the view point at time t_0 can be considered insufficient for a decision-maker who is intrinsically an intertemporal optimizer.

While a detailed analysis of the break even point is beyond the scope of this paper, we notice that t^* does not depend neither on wealth nor on risk aversion α . It depends only on the market price of risk δ , on the time horizon T and on initial time t_0 . Table 1 reports the value of t^* with some typical values of δ and T , when the initial time is $t_0 = 0$: $\delta = 0.1, 0.2, 0.3, 0.4, 0.5$ and $T = 10, 20, 30, 40$ (see Vigna, 2014).

The break even point increases with δ and with T . This is expected, because when $\delta = 0$ the three strategies collapse into the riskless strategy (where all the fund is invested continuously in the riskless asset) and there is no difference between portfolios, value functions and reward functions. The value and reward functions are trivially the same also in the degenerate case $T = t_0$. We notice that t^* varies between 0.16 years, i.e. two months for 10

Sharpe ratio δ	Time horizon T			
	10	20	30	40
0.1	0.16	0.64	1.42	2.49
0.2	0.62	2.31	4.82	7.96
0.3	1.28	4.32	8.31	12.79
0.4	1.99	6.05	10.77	15.69
0.5	2.65	7.28	12.23	17.23

Table 1: Break even point t^* with different Sharpe ratios δ and time horizons T (in years).

years and small δ ($\delta = 0.1$) and 17 years for 40 years and high δ ($\delta = 0.5$). In relative terms, t^* ranges between 1% and 6% of T for small δ , and between 26% and 43% of T for high δ .

The intuition behind the fact that t^* increases with δ and T can be the following. The dynamically optimal strategy is more aggressive than the Nash-equilibrium strategy by a factor $e^{\delta^2(T-t)}$ (see (8) and (9)). When δ or T increase, the dynamically optimal strategy becomes even more aggressive than the Nash-equilibrium strategy. A more aggressive strategy increases to a larger extent the variance of final wealth, and to a lower extent the mean of final wealth, hence the global effect is a reduction of the intertemporal reward, and this happens with higher severity with the dynamically optimal strategy than with the Nash-equilibrium strategy. Therefore, an increase in δ or T moves the break even point (that is the time point when the expected rewards of the two strategies switch) more far in the future, i.e., t^* increases.

Remark 5. The existence of a break even point between two strategies in a different context of time inconsistency has been found also in Balter, Mahayni & Schweizer (2019): in an optimal investment framework with ambiguity (where ambiguity is the source of time inconsistency) they compare the precommitment against the time-consistent Nash equilibrium strategies, and they find that there is a point of regret (that lies in the middle part of the investment horizon) after which the precommitted investor regrets his decision. In their case, the ratio t^* to T is in the range [37%, 71%].

3.2.3 Comparison of expected reward functions at time t : comparison among time-inconsistent and time-consistent strategies

We now intend to compare at future time t the precommitment time-inconsistent strategy with the two time-consistent strategies considered above. In order to do this, we need to calculate the expected value in t_0 of the reward function for the precommitment strategy (19):

$$R(t, u^p; t_0, x_0) = \mathbb{E}_{t_0, x_0} [J^p(t, X_t^p)] = \mathbb{E}_{t_0, x_0} [\mathbb{E}_{t, X_t^p}(X_T^p) - \alpha \mathbb{V}_{t, X_t^p}(X_T^p)] .$$

As we noticed in Section 3.2, and differently from the other two cases, in the precommitment case the value of the reward $J^p(t, X_t^p)$ *does not* coincide with the value function $V^p(t, X_t^p)$ calculated at time t , because the precommitment strategy is *not* time-consistent. Indeed, if the precommitted investor finds himself at time t with fund X_t^p , he will still apply the precommitted strategy (7). Therefore, the value of his reward $J^p(t, X_t^p)$ will *not* be the supremum of all possible values, the value function at (t, X_t^p) : the value function could be reached only by applying a *new* precommitment strategy with starting point t and initial wealth X_t^p . In other words, for the static precommitted investor the “value function” has meaning only at time t_0 and has no meaning after t_0 . However, inspired by (18), we can still assume that the reward for the precommitted investor at time t with wealth X_t^p is given by $J^p(t, X_t^p)$, and we can calculate its expectation at t_0 , $\mathbb{E}_{t_0, x_0} [J^p(t, X_t^p)]$.

If the static investor has wealth x_t^p at time $t > t_0$, and adopts the investment strategy (7), then his future wealth at time $\tau > t$ follows the dynamics given by the SDE:

$$\begin{cases} dX_\tau^p = X_\tau^p [r + u^p(\tau, X_\tau^p)(\lambda - r)] d\tau + X_\tau^p u^p(\tau, X_\tau^p) \sigma dW_\tau \\ X_t^p = x_t^p \end{cases} \quad (37)$$

where

$$u^p(\tau, x) = \frac{\delta}{\sigma} \frac{1}{x} (K e^{-r(T-\tau)} - x)$$

with

$$K = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}. \quad (38)$$

Let us define the new stochastic process

$$Z_\tau = Ke^{-r(T-t_0)} - X_\tau^p e^{-r(\tau-t_0)}, \quad (39)$$

with K given by (38). By applying Ito's lemma to Z_τ , we get its dynamics for $\tau > t$:

$$\begin{cases} dZ_\tau = -\delta^2 Z_\tau d\tau - \delta Z_\tau dW_\tau \\ Z_t = Ke^{-r(T-t_0)} - X_t^p e^{-r(t-t_0)}. \end{cases} \quad (40)$$

Therefore

$$Z_\tau = Z_t e^{-\frac{3}{2}\delta^2(\tau-t) - \delta(W_\tau - W_t)} \quad (41)$$

where Z_t is given by (40). Plugging (41) into (39), after some simplifications, we get the solution to (37), i.e. the dynamics of X_τ^p for $\tau > t$:

$$\begin{aligned} X_\tau^p = & \left[x_0 e^{r(\tau-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0) - r(T-\tau)} \right] + \\ & + \left[x_t^p - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0) - r(T-t)} \right] e^{(r-\frac{3}{2}\delta^2)(\tau-t)} \cdot e^{-\delta(W_\tau - W_t)}. \end{aligned} \quad (42)$$

Therefore

$$\begin{aligned} X_T^p = & \left[x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)} \right] + \\ & + \left[x_t^p - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0) - r(T-t)} \right] e^{(r-\frac{3}{2}\delta^2)(T-t)} \cdot e^{-\delta(W_T - W_t)}. \end{aligned} \quad (43)$$

Thus

$$\mathbb{E}_{t, x_t^p}(X_T^p) = \left(1 - e^{-\delta^2(T-t)}\right) \left(x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right) + x_t^p e^{(r-\delta^2)(T-t)}, \quad (44)$$

and

$$\mathbb{V}_{t, x_t^p}(X_T^p) = \left[x_t^p - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0) - r(T-t)} \right]^2 \left(e^{(2r-\delta^2)(T-t)} - e^{2(r-\delta^2)(T-t)} \right). \quad (45)$$

Therefore, the reward at time t for the precommitted investor is

$$\begin{aligned}
J^p(t, X_t^p) &= \mathbb{E}_{t, X_t^p}(X_T^p) - \alpha \mathbb{V}_{t, X_t^p}(X_T^p) = \\
&\quad \left(1 - e^{-\delta^2(T-t)}\right) \left(x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right) + X_t^p e^{(r-\delta^2)(T-t)} + \\
&\quad - \alpha \left[X_t^p - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right]^2 \left(e^{(2r-\delta^2)(T-t)} - e^{2(r-\delta^2)(T-t)}\right). \quad (46)
\end{aligned}$$

Taking expectation at time t_0 , we get:

$$\begin{aligned}
\mathbb{E}_{t_0, x_0}(J^p(t, X_t^p)) &= \mathbb{E}_{t_0, x_0}[\mathbb{E}_{t, X_t^p}(X_T^p) - \alpha \mathbb{V}_{t, X_t^p}(X_T^p)] = \\
&\quad \left(1 - e^{-\delta^2(T-t)}\right) \left(x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right) + \mathbb{E}_{t_0, x_0}(X_t^p) e^{(r-\delta^2)(T-t)} + \\
&\quad - \alpha \left[\mathbb{E}_{t_0, x_0}(X_t^p) - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right]^2 \left(e^{(2r-\delta^2)(T-t)} - e^{2(r-\delta^2)(T-t)}\right). \quad (47)
\end{aligned}$$

Using the dynamics (10), we get

$$\mathbb{E}_{t_0, x_0}(X_t^p) = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2-r)(T-t)} \left[e^{\delta^2(t-t_0)} - 1\right]. \quad (48)$$

Plugging (48) into (47), after some simplifications, we get the expected value in t_0 of the reward function at time t for the precommitted investor:

$$R(t, u^p; t_0, x_0) = \mathbb{E}_{t_0, x_0}(J^p(t, X_t^p)) = x_0 e^{r(T-t_0)} + \frac{1}{4\alpha} \left[2e^{\delta^2(T-t_0)} - e^{\delta^2(T-t)} - 1\right]. \quad (49)$$

Remark 6. Notice that using (49) to calculate $R(t, u^p; t_0, x_0)$ in t_0 and T , we get

$$R(t_0, u^p; t_0, x_0) = x_0 e^{r(T-t_0)} + \frac{1}{4\alpha} \left[e^{\delta^2(T-t_0)} - 1\right],$$

which, as expected, coincides with (14), and

$$R(T, u^p; t_0, x_0) = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} \left[e^{\delta^2(T-t_0)} - 1\right],$$

which, as expected, coincides with (25).

We are now ready to compare the static precommitment strategy with the time-consistent strategies. The following results hold:

Theorem 3.4. *If $t_0 < T$, $\delta \neq 0$, and $x_0 > 0$,*

$$R(t, u^p; t_0, x_0) > R(t, u^e; t_0, x_0) \quad \text{for all } t \in [t_0, T], \quad (50)$$

$$R(t, u^p; t_0, x_0) > R(t, u^d; t_0, x_0) \quad \text{for all } t \in [t_0, T], \quad (51)$$

and

$$R(T, u^p; t_0, x_0) = R(T, u^d; t_0, x_0). \quad (52)$$

Proof. By defining the function

$$\Delta R^{pe}(t) = R(t, u^p; t_0, x_0) - R(t, u^e; t_0, x_0), \quad (53)$$

claim (50) is equivalent to the strict positivity of $\Delta R^{pe}(t)$ over $[t_0, T]$. By plugging (49) and (32) into (53), we get:

$$\Delta R^{pe}(t) = \frac{1}{4\alpha} \left[2e^{\delta^2(T-t_0)} - e^{\delta^2(T-t)} - 1 - \delta^2(T+t-2t_0) \right]. \quad (54)$$

The first derivatives of $\Delta R^{pe}(t)$ is

$$\frac{d(\Delta R^{pe}(t))}{dt} = \frac{\delta^2}{4\alpha} \left(e^{\delta^2(T-t)} - 1 \right) > 0 \quad \text{for } t \in [t_0, T],$$

implying that $\Delta R^{pe}(t)$ is increasing over $[t_0, T]$. Note that $\Delta R^{pe}(t_0) > 0$ by Proposition 3.2. Then, claim (50) follows.

By defining the function

$$\Delta R^{pd}(t) = R(t, u^p; t_0, x_0) - R(t, u^d; t_0, x_0), \quad (55)$$

claim (51) is equivalent to the strict positivity of $\Delta R^{pd}(t)$ over $[t_0, T]$, and claim (52) is

equivalent to $\Delta R^{pd}(T) = 0$. By plugging (49) and (33) into (55), we get:

$$\Delta R^{pd}(t) = \frac{1}{8\alpha} \left[e^{\delta^2(T-t)} - 1 \right]^2. \quad (56)$$

Then, claims (51) and (52) follow easily. \square

From Theorem 3.4 we see that the static precommitment strategy provides a higher expected reward for any time $t \in [t_0, T]$ than the two time-consistent strategies, meaning that the static precommitment strategy is never inferior to the other strategies. Similarly to Remark 1, this is obvious and due to the fact that the criterion $R(t, u; t_0, x_0)$ illustrates the degree of happiness at time t of the mean-variance optimizer *as measured at time t_0* . No surprise that the precommitment strategy gives an outcome at least as good as the other strategies. From Theorem 3.4 we also see that the difference $\Delta R^{pe}(t)$ between precommitment and Nash-equilibrium strategies is *increasing* over $[t_0, T]$, and reaches its maximum in T : this means that the positive difference between expected rewards becomes larger when time passes. On the contrary, the difference $\Delta R^{pd}(t)$ between precommitment and dynamically optimal strategies is *decreasing* over $[t_0, T]$, and is null in T : this means that the positive difference between expected rewards becomes smaller when time passes, and disappears at terminal time T .

3.3 Additional considerations about time consistency

In the previous sections we have compared the three strategies using an intertemporal reward-driven criterion. It is worth adding a few considerations that can shed further light on the differences among the three approaches.

In their seminal paper in Proposition 5.1 Björk & Murgoci (2010) prove that to each time-inconsistent problem it is possible to associate a standard time-consistent problem such that (i) the optimal value function of the standard problem is equal to the equilibrium value function of the time-inconsistent problem; (ii) the optimal control law of the standard problem is equal to the equilibrium strategy of the time-inconsistent problem.

This remarkable result implies that, given a time-inconsistent problem, there exist utility functions $U^1(\cdot)$ and $U^2(\cdot)$ such that the Nash equilibrium strategy associated to the time-

inconsistent problem coincides with the optimal control law found via dynamic programming solution to the time-consistent problem

$$\sup_{u \in \mathcal{U}} \mathbb{E}_{t_0, x_0} \left[\int_{t_0}^T U^1(s, X_s, u_s) ds + U^2(X_T) \right]. \quad (57)$$

In the case of the mean-variance problem we have $U^1(x) = 0$ and $U^2(x) = -1/(2\alpha)e^{-2\alpha x}$ (see Basak & Chabakauri, 2010 and Vigna, 2017). Therefore, enforcing time consistency with the Nash-equilibrium approach results in the application of the optimal policy for the *alternative* time-consistent standard problem (57). According to Vigna (2017), the Nash-equilibrium investor is *not preferences-consistent* (for the precise definition of preferences consistency we refer the reader to Vigna, 2017).

We find a similar situation with the precommitment policy. Indeed, thanks to the embedding result of Zhou & Li (2000), it is possible to prove that a target $\gamma \in \mathbb{R}$ exists such that at any time $t \in [t_0, T]$ the precommitted investor plays the optimal strategy for the following target-based problem:

$$\inf_{u \in \mathcal{U}} \mathbb{E}_{t_0, x_0} [(X(T) - \gamma)^2]. \quad (58)$$

This means that the precommitment strategy coincides with the strategy of an investor with target-based preferences with target γ (see also Menoncin & Vigna, 2017 and Menoncin & Vigna, 2020). Therefore, similar to the Nash-equilibrium strategy, also the precommitment strategy turns out to be the solution of the alternative time-consistent standard problem (58).

Thus, both the Nash equilibrium policy and the precommitment strategy are the optimal policies associated to two time-consistent problems (respectively, Problems (57) and (58)) that are different from the original mean-variance problem. In this sense, both the Nash-equilibrium investor and the precommitted investor lack preferences consistency (see Vigna, 2017).

4 Concluding remarks

In this paper we have addressed a comparison between three common approaches to time inconsistency for the mean-variance portfolio selection problem, namely, precommitment approach, consistent planning approach and dynamical optimality approach. While the comparison at initial time t_0 is trivial, we have provided a comparison also at time $t > t_0$, by defining an intertemporal stochastic time- t -reward, based on the assumption that the individual does not change his mean-variance preferences and risk aversion over time. A comparison between the expected values at time t_0 of the stochastic time- t -reward shows that, standing at time t_0 , and expectedly, the precommitment approach beats the other two approaches, at the cost of being time-inconsistent. The comparison between the two time-consistent approaches, i.e., the consistent planning and the dynamical optimality approach, shows that the former dominates the latter from t_0 up to a unique break even point $t^* \in (t_0, T)$ and is dominated by the latter from t^* to T .

The main message that one can probably get from this analysis is that when there is a problem that gives rise to time inconsistency there is no clear-cut answer to the issue “what is the right thing to do”. A normative approach that pretends to be universal fails to provide convincing arguments, for the appropriate behaviour is dictated not only by the non-linear optimizing criterion but also by other subjective factors, such as the attitude towards time consistency, and the importance given to different time intervals and singular points in time. Instead, we consider a philosophical approach more appropriate.

References

- Bajeux-Besnainou, I. & Portait, R. (1998). Dynamic asset allocation in a mean-variance framework, *Management Science* **44**: S79–S95.
- Balter, A., Mahayni, A. & Schweizer, N. (2019). Time-consistency of optimal investment under smooth ambiguity, *Preprint. Available at SSRN:https://ssrn.com/abstract=3040545*.

- Basak, S. & Chabakauri, G. (2010). Dynamic mean-variance asset allocation, *Review of Financial Studies* **23**: 2970–3016.
- Bensoussan, A., Wong, K. & Yam, S. (2019). A paradox in time-consistency in the mean-variance problem?, *Finance and Stochastics* **23**: 173–207.
- Björk, T. (1998). *Arbitrage Theory in Continuous Time*, Oxford University Press.
- Björk, T., Khapko, M. & Murgoci, A. (2017). On time-inconsistent stochastic control in continuous time, *Finance and Stochastics* **21**: 331–360.
- Björk, T. & Murgoci, A. (2010). A general theory of markovian time inconsistent stochastic control problems, *Preprint, SSRN: <http://ssrn.com/abstract=1694759>*.
- Björk, T. & Murgoci, A. (2014). A theory of markovian time-inconsistent stochastic control in discrete time, *Finance and Stochastics* **18**: 545–592.
- Björk, T., Murgoci, A. & Zhou, X. (2014). Mean variance portfolio optimization with state dependent risk aversion, *Mathematical Finance* **24**: 1–24.
- Chen, S., Li, Z. & Zeng, Y. (2014). Optimal dividend strategies with time-inconsistent preferences, *Journal of Economic Dynamics and Control* **46**: 150–172.
- Christensen, S. & Lindensjö, K. (2018). On finding equilibrium stopping times for time-inconsistent markovian problems, *SIAM Journal on Control and Optimization* **56**: 4228–4255.
- Cong, F. & Oosterlee, C. (2016). On pre-commitment aspects of a time-consistent strategy for a mean-variance investor, *Journal of Economic Dynamics and Control* **70**: 178–193.
- Cui, X., Li, D. & Shi, Y. (2017). Self-coordination in time inconsistent stochastic decision problems: A planner-doer game framework, *Journal of Economic Dynamics and Control* **75**: 91–113.
- Cui, X., Li, D., Wang, S. & Zhu, S. (2012). Better than dynamic mean-variance: time inconsistency and free cash flow stream, *Mathematical Finance* **22**: 346–378.

- Czichowsky, C. (2013). Time consistent mean variance portfolio selection in discrete and continuous time, *Finance and Stochastics* **17**: 227–271.
- Karnam, C., Ma, J. & Zhang, J. (2017). Dynamic approaches for some time inconsistent problems, *The Annals of Applied Probability* **27**: 3435–3477.
- Li, D. & Ng, W.-L. (2000). Optimal dynamic portfolio selection: multiperiod mean-variance formulation, *Mathematical Finance* **10**(3): 387–406.
- Menoncin, F. & Vigna, E. (2017). Mean-variance target-based optimisation for defined contribution pension schemes in a stochastic framework, *Insurance: Mathematics and Economics* **76**: 172–184.
- Menoncin, F. & Vigna, E. (2020). Mean-variance dynamic optimality for DC pension schemes, *European Actuarial Journal* **10**: 125–148.
- Pedersen, J. & Peskir, G. (2017). Optimal mean-variance portfolio selection, *Mathematics and Financial Economics* **11**: 137–160.
- Pollak, R. A. (1968). Consistent planning, *The Review of Economic Studies* **35**: 201–208.
- Powell, W. (2011). *Approximate Dynamic Programming*, John Wiley and Sons.
- Richardson, H. (1989). A minimum variance result in continuous trading portfolio optimization, *Management Science* **35**: 1045–1055.
- Shi, Y., Li, X. & Cui, X. (2017). Better than pre-committed optimal mean-variance policy in a jump diffusion market, *Mathematical Methods of Operations Research* **85**: 327–347.
- Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization, *The Review of Economic Studies* **23**: 165–180.
- Van Staden, P. M., Dang, D.-M. & Forsyth, P. A. (2018). Time-consistent mean-variance portfolio optimization: A numerical impulse control approach, *Insurance: Mathematics and Economics* **83**: 9–28.

- Vigna, E. (2014). On efficiency of mean variance based portfolio selection in defined contribution pension schemes, *Quantitative Finance* **14**: 237–258.
- Vigna, E. (2017). Tail optimality and preferences consistency for stochastic optimal control problems, *Carlo Alberto Notebook* 502.
- Wang, J. & Forsyth, P. A. (2011). Continuous time mean variance asset allocation: A time-consistent strategy, *European Journal of Operational Research* **209**: 184–201.
- Wang, J. & Forsyth, P. A. (2012). Comparison of mean variance like strategies for optimal asset allocation problems, *International Journal of Theoretical and Applied Finance* **15(2)**.
- Zhou, X. & Li, D. (2000). Continuous-time mean-variance portfolio selection: A stochastic LQ framework, *Applied Mathematics and Optimization* **42**: 19–33.